DOCUMENT RESUME

ED 161 895 TM 007 381

AUTHOR Ory, John C.; Poggio, John P.
TITLE The Effect of Varying Response-Mode Formats on Responses to Affective Measures.
PUB DATE Mar 78
NOTE 21p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Toronto, Ontario, Canada, March, 1978)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Affective Tests; *Answer Sheets; Attitude Tests; *Cognitive Style; Higher Education; Personality Tests; *Response Mode; *Response Style (Tests); Secondary Education; Secondary School Students; *Testing; Test Items; Undergraduate Students
IDENTIFIERS Education Scale; Hidden Figures Test; Personality Research Form

ABSTRACT

The study investigated the effect of response mode variation (the use or non-use of separate answer sheets) on subject responses (N = 1017) to affective instruments treating both instrument and subject characteristics. Instrument characteristics examined included attitudinal versus personality scale composition, scale items written in first versus third-person mode and variations in response actions (circling, writing or darkening response). Subject characteristics investigated were individual differences due to cognitive style (flexibility of closure), and school grade level (junior high, high school or college). Results indicated significant score differences across scales with and without answer sheets. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED161895

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The Effect of Varying Response-Mode Formats on Responses to Affective Measures

John C. Ory
Measurement and Research Division
University of Illinois

John P. Poggio
Educational Psychology and Research
University of Kansas

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John C. Ory

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Paper presented at the Annual Meeting of the
National Council on Measurement in Education

Toronto, Canada

March 1978

Printed in U.S.A.

TM007 381

The Effect of Varying Response-Mode

Formats on Responses to Affective Measures

Research has indicated that response sets cause a person to respond differently to identical content items presented in varying forms (Huck and Jacko, 1973; Lederman, 1971; Mathews, 1930). Further, these same response sets tend to influence an individual's test performance and, in doing so, mask the test score so as to leave it ambiguous or uninterpretable (Cronbach, 1946). Due to the existence and masking effect of response sets, psychometricians have faced a need to detect, analyze and attempt to eliminate or control these sources of measurement error.

One area of test form variation has been virtually overlooked in the search for response set influences. This area involves the possible existence of response set behavior with affective tests presented under varying response-mode conditions, specifically, answer sheet or booklet responding. While research has consistently shown that the response modes of separate answer sheets and within booklet responding do not differentially effect examinee performance on cognitive assessment (Cashen and Ramseyer, 1969; Clark, 1968; Dunlap, 1940; Loper, 1939; Ramseyer and Cashen, 1971), the affective assessment has been virtually overlooked.

The intent of the study was to investigate the possible effect of response mode variation (the use or non-use of separate answer sheets) on subject responses to affective measures. By manipulating instrument and subject characteristics, an attempt was made to identify which kinds of affective instruments reflected a response-mode-set along with which individuals were most likely to be affected by such a behavioral bias.

Instrument Characteristics

To investigate the possible effects of an answer sheet response set with affective instruments, two categories of such devices were examined: attitudinal scales and personality inventories. The following measures were selected as being representative of each category.

The Education Scale, developed by Rundquist and Sletto (1963), was selected as an attitudinal measure. The scale assesses a respondent's attitude toward formal education.

High scores[1] indicate positive attitudes toward the value of education. The scale contains 22 items responded to on a Likert type scale from strongly agree to strongly disagree. The authors report a split-half reliability of .83, and a test-retest reliability over three weeks of .85. The instrument is characterized by items that are all stated in the form of third-person generalities. For example:

Savings spent on education are wisely invested.

The "Social Recognition" subscale of the Personality Research Form (PRF; Jackson, 1967) was chosen as being representative of many such well constructed personality inventories. The subscale contains 20 items that are all stated in the first-person mode. For example:

I consider it important to be held in high esteem by those I know.

High scores (see Footnote 1) are defined by Jackson (1967) as "concerned about reputation and what other people think of him; works for the approval and recognition of others" (p. 7). Split-half reliability of .91 and a stability coefficient over one week of .78 are reported in the instrument's manual.

Person Mode of Affective Items

After the selection of the representative scales, attention was directed to research by Getzels and Walsh (1958). Results of their investigation indicated differing response patterns when affective items were written in a first or third-person mode. In an attempt to control this differential occurrence and to examine the possible interaction between item mode and the hypothesized answer sheet response set, equivalent forms were produced for each scale, one form written in the first person mode and the other form written in the third-person condition. For example:

Third-Person Original: *People are foolish to keep going to school if they can get a job.*

First-Person Alteration: *If I can get a job I would be foolish to keep going to school.*

Response Action

In addition to the manipulation of person mode, another instrument characteristic

examined was the response action (e.g., circling, darkening) required of the examinee. To determine whether or not there existed an interaction between response action and a response-mode-set, such that the behavioral set occurred only for particular response actions, three actions were examined. The three response actions included circling, writing or darkening a response choice whether on a separate answer sheet or within a test booklet. These three actions are used with many of today's standardized or non-standardized classroom and research affective instruments.

Therefore, by manipulating the type of affective instrument (attitudinal vs. personality), the person mode of the items (1st vs. 3rd person) and the response action required by the instrument (circling, writing or darkening), the study attempted to identify which, if any, instrument characteristics enhanced a person's susceptibility to respond differently to affective items presented with and without a separate answer sheet.

Subject Characteristics

Should a response set exist that is attributed to the use or non-use of separate answer sheets, does the response set effect all individuals? Are some individuals more inclined to respond differently with and without answer sheets than others? To begin answering these inquiries two subject characteristics were examined: school grade level and cognitive style.

Grade Level

Students are exposed to an increasing number and variety of response modes (e.g., answer sheets, punch cards) as they progress through grade levels. Differences therefore exist in the number and kind of experiences that students within different grade levels have with certain response modes. Due to these differences in response mode familiarity (or lack thereof) it was hypothesized that the grade level of the examinee may interact with his/her susceptibility to a response-mode-set. This interaction would be reflected as grade level differences in the magnitude of response set behavior.

Cognitive Style

Cognitive style was chosen as the other subject characteristic to be examined. During the period of time between the reading of and then the responding to an affective item an examinee determines his/her response. An understanding of response processes during this interval of time is important to the understanding of the responses. Consequently, one such area of response processing involves individual differences in analytical functioning or cognitive style. A specific cognitive style that is relevant to the question of subject characteristics and response set susceptibility is identified as flexibility of closure or field dependence-independence (Witkin, 1954). This style of processing involves the ability to keep in mind a definite configuration so as to identify it in spite of perceptual distractions.

It was an hypothesis of this study that the existence of the answer sheet response set may be due in part to the distraction caused by leaving the perceptual field and responding on a separate answer sheet (outside of the field). Individual differences in the susceptibility to such "field" distraction, (level of field dependence) may account for some individuals being more effected by the answer sheet response set than others.

Methodology

To investigate the existence of the hypothesized answer sheet response set and its possible interaction with selected instrument and subject characteristics, 10 experimental test forms were constructed. The 10 forms were divided into two sets of five forms each. One set included forms with items from both the attitude and the personality instruments written in the first-person mode, while the second set included forms with items from both instruments written in third-person generalities. Within each set the five forms were constructed with the following response formats:

Form 1: Subjects were directed to respond on a separate answer sheet prepared by the investigators in a format in which the item numbers are listed along with codes SA, A, MA, MD, D and SD. Subjects were instructed to circle their choice response according to whether they strongly agree, agree, mildly agree, etc.

Form 2: A horizontal line of abbreviations SA, A, MA, MD, D and SD was placed under each item presented in the test booklet. Subjects were directed to respond to each item by circling their choice responses, according to whether they strongly agree, agree, mildly agree, etc.

Form 3: Subjects were directed to respond on a separate answer sheet prepared by the investigator. Item numbers were followed by a blank line on which respondents were to specify their response by writing SA, A, MA, MD, D or SD depending on their response choice of strongly agree, agree, mildly agree, etc.

Form 4: A blank line was presented to the immediate left of each item in the test booklet in which respondents were directed to write the abbreviated code, SA, A, MA, MD, D or SD to signify their choice of strongly agree, agree, mildly agree, etc.

Form 5: Subjects were directed to respond on a separate answer sheet. The answer sheet was a standard IBM response sheet, which was responded to by blackening the response choice on the sheet.

Also included on each experimental form were twenty "filler" items. These items were written in a person mode consistent with the other items on the form and were included to mask the intent of the attitudinal and personality scales.

Subjects included 1,017 students across three levels of education -- 405 junior high school students (7th and 8th grades), 294 high school students (juniors and seniors) and 318 college students (juniors and seniors). The subjects were randomly administered one of the ten experimental forms and the Hidden Figures Test (Witkin, 1954). The latter provided a measure of the subjects' level of closure flexibility. Following this first testing, subjects were readministered another experimental form three weeks after the initial testing under an identical response mode condition (answer sheet/circling, answer sheet/writing, booklet/circling, booklet/writing and answer sheet/blackening) but with items written in the alternate person mode. For example, if a subject had initially responded to affective items written in a first person mode by circling a response choice on a separate answer sheet, he/she was asked three weeks

later to respond to the same items written in a third-person mode by circling a response choice on a separate answer sheet.

Data Analysis

For each subject the scale scores on the measure of attitude and the personality inventory were computed and used as the dependent variable measures. A repeated factoral multivariate analysis of variance was computed examining these scale scores. The tested design included the repeated dimension of person mode (1st or 3rd person) and the five response mode conditions (answer sheet/circling, booklet/circling, answer sheet/writing, booklet/writing and answer sheet/darkening) crossed with the independent variables of grade level (junior high, high school or college) and the subjects' level of closure flexibility (high, middle, or low 1/3 in each age referent group). All hypotheses were tested at the .05 level of significance. Univariate ANOVA statistics and simple effects analyses were used as post hoc procedures as necessary.

Results

The source table for the MANOVA is presented in Table 1. Results of the analysis indicated three statistically significant ($p < .05$) main effects of Grade (junior high, high school or college), Response Mode (answer sheet/circling, booklet/circling, answer sheet/writing, booklet/writing, answer sheet/darkening) and Person (1st or 3rd person mode) and three significant interaction effects of HFT (closure flexibility) by Form, Person by Grade and Person by Grade by HFT by Response Mode. Univariate ANOVA statistics were computed to further analyze the statistically significant main and interaction effects. Results of these univariate ANOVA statistics for each dependent variable are presented in Tables 2 and 3.

- - - - - - - - - - - - - - - - - - - -
INSERT TABLES 1, 2 and 3 ABOUT HERE
- - - - - - - - - - - - - - - - - - - -

ANOVA 1: Attitude (Education) Scale Scores

Due to the confounding contributions of the significant main effects to the two higher order interactions, interpretation of the ANOVA statistic for the Education scale

scores is limited to discussion of the interaction effects. The statistically significant ($p < .01$) interaction effect of Person by Grade indicated that Education scale scores were differentially affected across grade levels by the presentation of items in first or third person. To isolate these differential effects, procedures described by Marascuilo and Levin (1970) were used to test estimated interaction components. Results indicated that only the college subject scores were significantly affected by the person mode of the items. A difference of approximately five scale points was found between college mean responses to first and third person Education scale scores (52.95 vs. 58.14), with first person items producing a higher positive attitude towards education.

The second statistically significant ($p < .01$) interaction effect was that of HFT by Response Mode. The significance of this interaction effect indicated that the Education scale scores of individuals with varying ability on the closure flexibility factor were differentially affected by the style of response mode presented. Again, post hoc procedures described by Marascuilo and Levin were followed to identify these effects. Results revealed that a single estimated interaction component was statistically significant ($p < .05$). The individual interaction component of Form Five (IBM answer sheet-darkening) and the lower one-third closure flexibility group provided a significant contribution to the interaction. Thus, of the students tested, those with relatively low closure flexibility (higher field dependence) scored significantly higher on the Education scale using the IBM answer sheet than did all other students. The significantly higher mean score for these field dependent - IBM form respondents reflected a less positive attitude toward education. A graph of the interaction component means is presented in Figure 1.

- - - - - - - - - - - - - - -
Insert Figure 1 About Here
- - - - - - - - - - - - - - -

To further investigate the significant interaction effect of HFT by Response Mode, hypothesized contrasts between cell interaction effects were computed. Examined contrasts across levels of closure flexibility included in-booklet versus out-of-

booklet responses, circled versus written responses and the IBM answer sheet responses versus a combination of all other form responses. Contrast results indicated a statistically significant ($p < .05$) difference between mean responses recorded in-booklet (Forms 2 + 4) and out-of-booklet (Forms 1 + 3 + 5) for students with relatively low ability on the closure flexibility factor. Those students identified as field dependent individuals were most affected by in- and out-of-booklet responding, reflecting a higher positive attitude toward education within the test booklet ($\bar{X} = 49.92$) than on separate answer sheets ($\bar{X} = 53.57$).

To summarize, results of the statistical analyses on Education scale scores indicated a significantly higher positive regard for education made by college students when education scale items were written in first person rather than third person. Also, individuals reflecting relatively low closure flexibility (field dependent individuals) responded significantly higher on Form 5 (IBM answer sheet) than did all other subjects, whereas Form 2 demonstrated the most stable responses across levels of closure flexibility. This same low ability group responded differently when responding in-booklet as compared to out-of-booklet. Their in-booklet scores reflected a significantly higher positive attitude toward education than did their out-of-booklet responses.

ANOVA 2: Personality Scale (Social Recognition) Scores

Results of the analysis of variance computed on Social Recognition scale scores revealed statistically significant ($p < .05$) main effects of Grade and Person, and a significant interaction of Person by Grade by HFT by Response Mode.

To investigate the four-way interaction, post hoc ANOVA statistics, one for each grade level data set, were computed. Due to the simple effects investigative purposes of these analyses, only the statistical significance of the highest order interaction (Person by HFT by Response Mode) was considered. This three-way interaction effect of Person by HFT by Response Mode was statistically significant ($p < .01$) for the college level data set, approached significance ($p = .055$) in the high school group, and failed to reach statistical significance in the junior high data set

($p = .46$). Results indicated that the dimensions of item person mode, level of closure felxibility and response mode presentation tended to affect the student's Social Recognition scores similarly yet significantly more as they reached the higher grade levels. A consideration of the significant college group interaction contributes to an interpretation of the original four-way interaction.

The three graphs presented in Figure 2, one for each level of closure flexibility within the college age group, depict this three-way interaction. Visual comparison of the plotted Social Recognition scale means indicated that there was greater score

- - - - - - - - - - - - - - -
Insert Figure 2 About Here
- - - - - - - - - - - - - - -

variability across first and third person response forms for increased levels of field dependence. The low HFT ability group (field dependents) recorded more score variability across forms that did the middle HFT ability group, which demonstrated greater score variability across forms than did the high HFT ability group (field independents). Thus, variation in response mode appeared to have the greatest effect on the range of mean scores across person modes for individuals who were more field dependent than field independent or for individuals more influenced by the response mode "field." Similar behavioral patterns were revealed through collapsed comparisons of in-booklet and out-of-booklet scale scores across levels of closure flexibility. The disparity between means for responses recorded on a separate answer sheet (Forms 1 + 3 + 5) versus those in the test booklet (Forms 2 + 4) was statistically significantly ($p < .05$) greater for the low field dependent group versus the middle field dependent group (3.02 vs. -1.33) and for the low field dependent group contrasted with the high field independent group (3.02 vs. -1.14).

In summary, the effects on Social Recognition scale scores attributed to Person, HFT and Response Mode sources of variation were similar across grades but significant only in the college data set. Within the college grade level, field dependent students were more affected by the variation of the five response modes than were field independent students. This same pattern was observed for the overall comparison of in- versus out-of-booklet responses.

Discussion and Conclusions

Analyses of the Education and Social Recognition scale scores revealed somewhat similar results. As predicted by hypotheses of the study, variation in response mode presentation interacted with the cognitive style and age of respondent to reveal significant score differences across response forms. These differences were found for both attitudinal and personality scale scores. Analyses indicated that field dependent students demonstrated a greater range of scores across the response forms and between in- and out-of-booklet responding than did field independents. Results also indicated that changes in response mode affected scale scores significantly more in the higher grade levels. There was, however, no consistent relationship between a scale's susceptibility to the response-mode-set and the scale's item person mode. Overall, results indicated that differences in response mode presentation interacted with the cognitive style and age of respondent to reveal significant score differences across response forms. Specifically, field dependent participants, as they grow older, demonstrated significant differences in their responses recorded on separate answer sheets versus within the test booklet.

From the results of the study we have identified a response-mode-set which acted as a significant source of score variation in the affective scales investigated. While the effects of such a response set appeared through the interactions of response mode variation, cognitive style and age of respondent, it is believed that these effects would have been detected outside the theoretical framework of interactive variables studied. However, the significant interaction has served to present evidence of the construct validity of such a response-mode-set. Through the interaction of form and trait, a rationale (individual differences in cognitive style and level of education) for the existence of the response-mode-set could be provided. Of course, the conclusions drawn are based on subject responses to particular attitude (Education) and personality (Social Recognition) instruments. In order to generalize results to affective scales in general, additional instruments must be examined across different populations.

The significance of study results is the identification of sources of variation which influence affective test scores other than those attributed to the latent trait of the scale's content. Potentially, these sources of variation contribute to measurement error, and result in a confounding affective score interpretation. Given the results of this investigation both developers and users of affective instruments need to be aware of the possible effect that response mode alteration may have on the psychometric properties of reliability, validity and item parameters. Of practical significance to instrument developers is the impact of such measurement error on the standardization process of test construction. If a standardized test was to provide alternative modes for examinee responding, yet only provide a single set of norms developed over all mode responses, results of this study would indicate that the norms may be inaccurate. Instead, developers of affective measures utilizing multiple responding modes (e.g., MMPI) should investigate possible effects due to a response-mode-set as well as construct separate norms for each response mode used. For similar reasons, users of affective instruments should be more attentive to a scale's normed or original response mode before altering it for the convenience of their particular use.

In conclusion, research findings indicated the existence of a response-mode-set which acted as a source of score variation within selected affective self-report measures. Further research is needed to better understand this source of score variation, analyze its consequences and provide ways for its control or elimination. Future research may utilize the multitrait-multimethod matrix technique of Campbell and Fiske (1959) to assess the impact of response-mode alteration on affective scale validity. Scales utilizing varying response modes but assessing the same psychological construct should exhibit convergent validity, whereas scales using similar response modes but assessing disimilar psychological constructs should exhibit discriminant validity. Differences in these anticipated patterns may be due in part to a response-mode-set.

References

Campbell, D. T. & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

Cashen, V. M. & Ramseyer, G. C. The use of separate answer sheets by primary age children. Journal of Educational Measurement, 1969, 6, 155-158.

Clark, Carl A. The use of separate answer sheets in testing slow-learning pupils. Journal of Educational Measurement, 1968, 5, 61-64.

Cronbach, L. J. Response sets and test validity. Educational and Psychological Measurement, 1946, 6, 475-495.

Dunlap, J. W. Problems arising from the use of a separate answer sheet. Journal of Psychology, 1940, 10, 2-48.

Getzels, J. W. & Walsh, J. J. The method of paired direct and projective questionnaires in the study of attitude structure and socialization. Psychological Monographs, 1958, 72.

Huck, S. W. & Jacko, E. J. Effect of varying response format of the Alpert-Haber Achievement Anxiety Test. Journal of Counseling Psychology, 1973, 21, 159-163.

Jackson, D. N. Personality research form. New York: Research Psychologists Press, Inc., 1967.

Lederman, E. Effect of the response format on the differential measurement of traits in the Thorndike Dimensions of Temperament Personality Inventory. ERIC, ED 052 249, 1971.

Loper, J. F. A study of the use of the separate answer sheet at the third and fifth grade levels. Unpublished master's thesis, State University of Iowa, 1939.

Marascuilo, L. A. & Levin, J. R. Appropriate post hoc comparisons for interaction and nested hypothesis in analysis of variance designs: The elimination of Type IV errors. American Educational Research Journal, 1970, 1, 397-421.

Mathews, C. O. The effect of the order of printed response on an interest questionnaire. Journal of Educational Psychology, 1930, 20, 128-134.

Footnotes

[1]Scoring was reversed in the study such that lower rather than higher scores reflected "more" of the attitude or trait being measured.

Table 1

MANOVA Analysis for Education

and Social Recognition Scale Scores

Source of Variation	df		F-Ratio
Grade	4	1942	60.59**
HFT	4	1942	1.49
Form	8	1942	2.17*
Grade x HFT	8	1942	.78
Grade x Form	16	1942	1.20
HFT x Form	16	1942	2.27**
HFT x Form x Grade	32	1942	1.04
Person	2	971	67.49**
Person x Grade	4	1942	10.77**
Person x HFT	4	1942	.73
Person x Form	8	1942	.42
Person x Grade x HFT	8	1942	1.41
Person x Grade x Form	16	1942	1.63
Person x HFT x Form	16	1942	1.36
Person x Grade x HFT x Form	32	1942	1.65**

*$p \leq .05$

**$p \leq .01$

Table 2

Education Scale Score ANOVA Source Table

Source of Variation	df	MS	F-Ratio
Grade	2	15695.21	35.08**
HFT	2	1181.04	2.64
Form	4	1393.31	3.12*
Grade x HFT	4	175.47	.39
Grade x Form	8	419.38	.94
HFT x Form	8	1524.32	3.41**
Grade x HFT x Form	16	236.09	.53
Error	972	447.36	
Person	1	8303.67	127.10**
Person x Grade	2	1320.44	20.21**
Person x HFT	2	.97	.02
Person x Form	4	51.32	.79
Person x Grade x HFT	4	127.62	1.95
Person x Grade x Form	8	97.81	1.50
Person x HFT x Form	8	74.66	1.14
Person x Grade x HFT x Form	16	104.70	1.60
Error	972	65.33	

*$p \leq .05$

**$p \leq .01$

Table 3

Social Recognition Scale Score

ANOVA Source Table

Source of Variation	df	MS	F-Ratio
Grade	2	35831.45	109.58*
HFT	2	21.26	.07
Form	4	522.68	1.60
Grade x HFT	4	370.48	1.13
Grade x Form	8	548.12	1.68
HFT x Form	8	400.20	1.22
Grade x HFT x Form	16	450.49	1.38
Error	972	326.99	
Person	1	1079.95	15.37*
Person x Grade	2	47.85	.68
Person x HFT	2	98.66	1.41
Person x Form	4	2.79	.04
Person x Grade x HFT	4	63.61	.91
Person x Grade x Form	8	118.57	1.69
Person x HFT x Form	8	105.64	1.50
Person x Grade x HFT x Form	16	124.39	1.77**
Error	972	70.25	

*$p \leq .01$

**$p \leq .05$

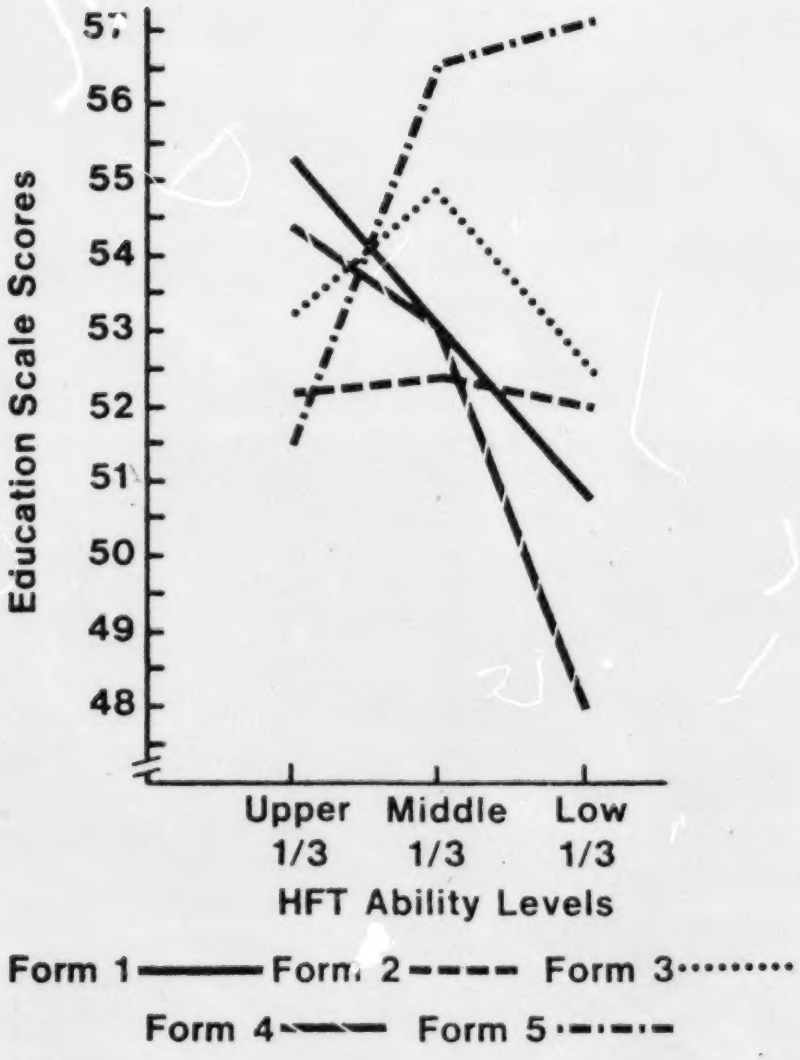

FIGURE 1. PLOT OF HFT BY RESPONSE FORM INTERACTION FOR EDUCATION SCALE SCORES.

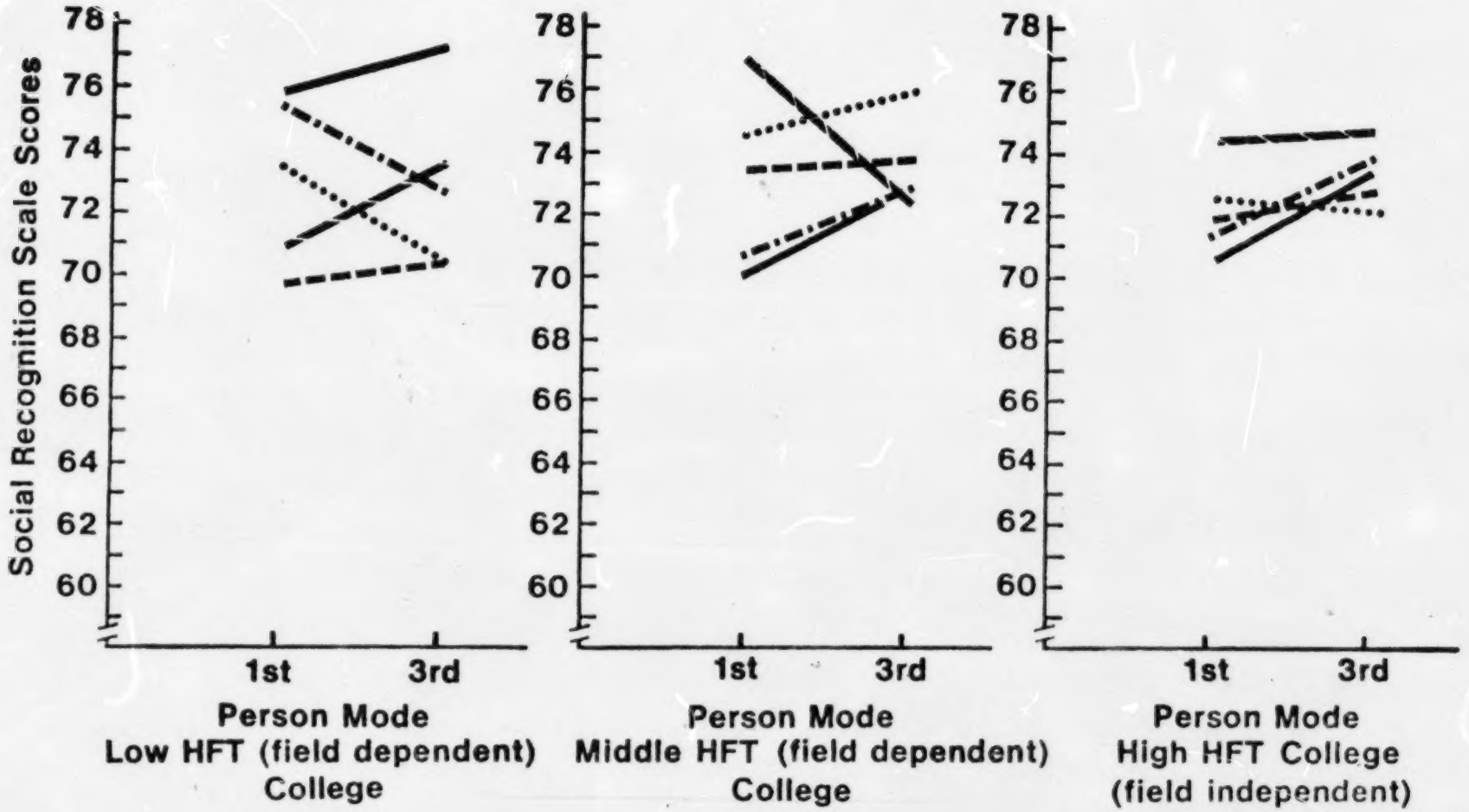

FIGURE 2. PERSON BY FORM INTERACTION EFFECTS BY LEVELS OF CLOSURE FLEXIBILITY FOR COLLEGE GROUP.